Fletcher and Dudbridge *BMC Medicine* 2014, **12**:195
http://www.biomedcentral.com/1741-7015/12/195

Spotlight on breast cancer

BMC Medicine

COMMENTARY

Open Access

Candidate gene-environment interactions in breast cancer

Olivia Fletcher[1,2*] and Frank Dudbridge[3]

Abstract

Gene-environment interactions have the potential to shed light on biological processes leading to disease, identify individuals for whom risk factors are most relevant, and improve the accuracy of epidemiological risk models. We review the progress that has been made in investigating gene-environment interactions in the field of breast cancer. Although several large-scale analyses have been carried out, only a few significant interactions have been reported. One of these, an interaction between *CASP8*-rs1045485 and alcohol consumption has been replicated, but others have not, including *LSP1*- rs3817198 and parity, and 1p11.2-rs11249433 and ever being parous. False positive interactions may arise if the gene and environment are correlated and the causal variant is less frequent than the tag SNP. We conclude that while much progress has been made in this area it is still too soon to tell whether gene-environment interactions will fulfil their promise. Before we can make this assessment we will need to replicate (or refute) the reported interactions, identify the causal variants that underlie tag-SNP associations and validate the next generation of epidemiological risk models.

Keywords: Gene-environment interaction, Breast cancer, Single nucleotide polymorphism

Background

Epidemiological studies have provided consistent evidence of associations between environmental (predominantly lifestyle and reproductive) factors and subsequent risk of breast cancer (BC). More recently, genome-wide association studies (GWAS) have identified more than 70 single nucleotide polymorphisms (SNPs) that influence breast cancer risk [1]. Detecting a gene-environment (GxE) interaction between a SNP and an environmental risk factor has the potential to shed light on the biological process leading to disease, identify women for whom these risk factors are most relevant, and improve the accuracy of epidemiological risk models [2]. A comprehensive review summarising the rationale for and the challenges of studying GxE interactions advocated a range of measures including supporting new and larger prospective studies, the reporting of stratified analyses as supplementary material and pre-planned analyses coordinated across multiple studies [2]. In this commentary we review progress in investigating GxE interactions in the field of BC. We define GxE interaction as the modification of the effect of a genetic risk factor by an environmental factor, assessed statistically by testing the effects of gene and environment for departure from additivity, on an appropriate scale (usually the log or logit in disease studies). We focus on GxE interactions between common SNPs and established risk factors for BC (Table 1), discuss the implications of testing marker SNPs rather than the underlying causal variants that they tag and consider whether GxE studies have fulfilled their potential for illuminating disease processes or predicting risk.

GxE interactions between previously reported SNPs and established risk factors for BC

The first large (that is, at least 5,000 cases and 5,000 controls) GxE study of this type was carried out within the Million Women Study [3]. In this analysis of 7,610 cases and 10,196 controls investigating potential GxE interactions between 12 SNPs and 10 established risk factors for BC there were no GxE interactions that were significant after adjusting for multiple testing. The most significant GxE interaction was between *CASP8*-rs1045485 and alcohol consumption (unadjusted $P = 0.003$). Since the

* Correspondence: Olivia.Fletcher@icr.ac.uk
[1]Breakthrough Breast Cancer Research Centre, The Institute of Cancer Research, London SW3 6JB, UK
[2]Division of Breast Cancer Research, The Institute of Cancer Research, London SW3 6JB, UK
Full list of author information is available at the end of the article

Table 1 Established risk factors assessed in GxE interaction studies

Established risk factor	Travis *et al.* [3]	Milne *et al.* [4]	Campa *et al.* [6]	Nickels *et al.* [5]	Barrdahl *et al.* [7]	Schoeps *et al.* [12]
Age at menarche (years)	X	X	X	X	X	X
Age at first birth (years)	X	X		X[a]	X	
Parous (% and/or number of live births)	X	X	X	X[a]	X	X
Breast fed (% in parous women)	X			X		
Menopausal status (% post-menopausal)	X					
Age at natural menopause (years)	X		X		X	
Use of oral contraceptives (yes/no or duration)				X	X	X
Use of HRT (yes/no or duration)	X		X	X		X[b]
Body Mass Index (kg/m^2)	X	X	X	X	X	X[c]
Height (m)	X		X	X	X	X
Alcohol consumption (g per day)	X		X	X	X	X
Smoking (pack-years)			X	X	X	
Family history of BC			X		X	X
Physical activity (hours per week)				X		

[a]The ten established risk factors reported by Nickels *et al.* (Table 2) counts parity and age at first live birth as a single factor; [b]the 10 established risk factors reported by Schoeps *et al.* (Table 2) counts combined estrogen-progesterone and estrogen only post-menopausal hormone therapy as two factors; [c]body mass index in pre- and post-menopausal women as two factors. BC, breast cancer; HRT, hormone replacement therapy.

publication of this report, there have been four further analyses of this type (Table 2), two from the Breast Cancer Association Consortium (BCAC) [4,5] and two from the Breast and Prostate Cancer Cohort Consortium (BPC3) [6,7]. Only one of these, the largest (23 SNPs in 34,793 cases and 41,099 controls) [5], reported statistically significant GxE interactions, namely between *LSP1*-rs3817198 and parity (number of live births), *CASP8*-rs1045485 and alcohol consumption (replicating the most significant finding in the Million Women study [3]) and 1p11.2-rs11249433 and ever being parous. However, none of these interactions was replicated in the largest BPC3 study (39 SNPs in 16,285 BC cases and 19,376 controls [7]). A meta-analysis of the BCAC and BPC3 data suggested a possible interaction between *SLC4A7*-rs4973768 and smoking status but replication of this result has not yet been attempted.

The Shanghai Breast Cancer Genetics Study tested for interactions using a risk score formed as the weighted sum of genotypes from 10 SNPs [8]. This would improve the power to detect a risk factor that has interactions with numerous SNPs, when there is insufficient power for the individual interactions. Although this study found no interactions with the risk score, this approach holds promise for identifying interacting risk factors in limited sample sizes.

Table 2 Details of GxE interaction studies comprising at least 5,000 cases and 5,000 controls

Reference	Study	SNPs	ERFs	Cases	Controls	Strongest interaction	Interaction effect size	Unadjusted *P*
Travis *et al.* [3]	Million women	12	10	7,610	10,196	*CASP8*-rs1045485; alcohol consumption[a]	1.24	0.003
Milne *et al.* [4]	BCAC (1)	12	4	26,349	32,208	*LSP1*- rs3817198; parity	1.05	0.002
Campa *et al.* [6]	BPC3 (1)	17	9	8,576	11,892	5p12-rs10941679; use of estrogen only HRT	1.22	0.0072
Nickels *et al.* [5]	BCAC (2)	23	10	34,793	41,099	*LSP1*- rs3817198; parity	1.06	2.4×10^{-6}
						1p11.2-rs11249433; ever parous	1.16	5.3×10^{-5}
						CASP8-rs17468277; alcohol consumption[a]	1.59	3.1×10^{-4}
Barrdahl *et al.* [7]	BPC3 (2)	39	10	16,285	19,376	6q25-rs2046210; alcohol consumption[a]	1.11	0.002
Schoeps *et al.* [12]	BCAC	71,527	10	34,475	34,786	21q22.12-rs10483028 and rs2242714; postmenopausal BMI	0.84	3.2×10^{-5}

[a]Alcohol consumption was defined as per additional 10 g in Travis *et al.* and Barrdahl *et al.* and as < or ≥20 g/per day by Nickels *et al.* BCAC, Breast Cancer Association Consortium; BMI, body mass index; BPC3, Breast and Prostate Cancer Cohort Consortium; ERFs, established risk factors for breast cancer; HRT, hormone replacement therapy.

Identification of novel risk SNPs through GxE interactions

SNPs with strong interaction effects may only be detectable when analysing gene and environment together, so they are missed by studies that consider SNPs in isolation. Methods that model and test the main and interaction effects of gene and environment jointly [9], or exploit the power of a case-only design while retaining robustness to possible gene environment dependence [10,11] have been developed for these purposes. Recently, several of these methods were applied to 71,527 SNPs with suggestive association with BC [12]. Interactions were identified between two SNPs on 21q22.12 (rs10483028 and rs2242714) and adult body mass index (BMI), and one in *ARID1B* (rs12197388) with age at menarche and with parity. rs12197388 was only significant in the joint test of main and interaction effects, and the interaction term was not significant but the two SNPs on 21q22.12 were detected via their interactions, and further studies of this nature may discover more interactions using these novel methods.

Using tag-SNPs as proxies for an underlying causal variant

The GxE studies described above have relied on using marker SNPs, predominantly identified through GWAS, as proxies for the underlying causal variants. This usually leads to a loss in power to detect interactions [13]. However, if gene and environment are dependent, a marker SNP can show an interaction even if there is no interaction at the causal variant [14]. These 'spurious interactions' tend to arise when the causal variant is rare in comparison to the marker. This may not often be the case, but it nevertheless warrants caution when reporting GxE interactions. We recently studied a marker SNP (rs10235235) associated with a reduction in urinary levels of an estrogen metabolite [15]. In 47,346 cases and 47,569 controls in the Collaborative Oncological Gene-environment Study (COGS) [1,16] this SNP showed (1) association with BC risk, (2) association with age at menarche in controls (but not cases) and (3) an interaction in which age at menarche modified the effect of rs10235235 on BC risk. In this example of a GxE interaction, therefore, the genetic risk factor (rs10235235) is dependent on the environmental risk factor (age at menarche), which could lead to a false positive [14]. Of the interactions reported to date, gene-environment dependence has been observed between *LSP1*-rs3817198 and parity and 21q22.12-rs10483208/rs2242714 and BMI. In cases such as these, an interaction can only be definitively established when all variation in the associated regions has been identified and tested.

Conclusions

Several of the recommendations made by Hunter in 2005 [2] have been pursued: large new prospective studies continue to be supported (for example the Breakthrough Generations study, a long-term cohort study focused on BC has recruited 112,049 women over the period 2003 to 2011 [17]), consortia of case–control (BCAC) and cohort studies (BPC3) have coordinated their efforts for analyses of data from >70,000 women and the results of stratified analyses have been conscientiously reported in supplementary tables [5,7]. However, one of the lessons of the first generation of BC GWAS [18-20] was that the per-allele disease odds ratios (ORs) associated with individual tag-SNPs were much smaller than hypothesised (1.07 to 1.26). Results from the first generation of GxE analyses suggest that the same may be true for interactions, with the reported interaction ORs ranging from 1.06 to 1.59. If marginal ORs of 1.07 to 1.26 require scans of several thousand cases and several thousand controls then, depending on the number of GxE interactions being tested, only GxE studies that include tens of thousands of cases and controls will have the power required to detect interactions. It is hardly a coincidence that the first study to report statistically significant GxE interactions was the first study of this order of magnitude [5]. Of the three significant interactions reported by Nickels and colleagues there is replication only for *CASP8*-rs1045485 and alcohol consumption. It is currently too soon to tell whether GxE interactions will shed light on disease processes and improve the accuracy of epidemiological risk models. Before we can make this assessment we will need to replicate or refute the reported interactions, identify the causal variants that underlie tag-SNP associations and validate the next generation of epidemiological risk models.

Abbreviations

BC: breast cancer; BCAC: the Breast Cancer Association Consortium; BMI: body mass index; BPC3: the Breast and Prostate Cancer Cohort Consortium; COGS: Collaborative Oncological Gene-environment Study; ERF: established risk factor; GWAS: genome-wide association study; GxE interaction: gene-environment interaction; HRT: hormone replacement therapy; OR: odds ratio; SNP: single nucleotide polymorphism.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

OF and FD wrote this commentary jointly. Both authors approved the final version.

Authors' information

OF is a group leader in genetic epidemiology at the Breakthrough Breast Cancer Research Centre. FD is professor of statistical genetics at the London School of Hygiene and Tropical Medicine.

Acknowledgements

We acknowledge support from Breakthrough Breast Cancer (OF) and the MRC (G1000718 and K006215; FD). The funding bodies had no role in the preparation of this manuscript.

Author details
[1]Breakthrough Breast Cancer Research Centre, The Institute of Cancer Research, London SW3 6JB, UK. [2]Division of Breast Cancer Research, The Institute of Cancer Research, London SW3 6JB, UK. [3]Department of Non-communicable Disease Epidemiology, London School of Hygiene and Tropical Medicine, London WC1E 7HT, UK.

Received: 16 September 2014 Accepted: 30 September 2014
Published online: 17 October 2014

References

1. Michailidou K, Hall P, Gonzalez-Neira A, Ghoussaini M, Dennis J, Milne RL, Schmidt MK, Chang-Claude J, Bojesen SE, Bolla MK, Wang Q, Dicks E, Lee A, Turnbull C, Rahman N, Breast and Ovarian Cancer Susceptibility Collaboration, Fletcher O, Peto J, Gibson L, Dos Santos Silva I, Nevanlinna H, Muranen TA, Aittomäki K, Blomqvist C, Czene K, Irwanto A, Liu J, Waisfisz Q, Meijers-Heijboer H, Adank M, *et al*: **Large-scale genotyping identifies 41 new loci associated with breast cancer risk.** *Nat Genet* 2013, **45:**353–361. 361e351-352.
2. Hunter DJ: **Gene-environment interactions in human diseases.** *Nat Rev Genet* 2005, **6:**287–298.
3. Travis RC, Reeves GK, Green J, Bull D, Tipper SJ, Baker K, Beral V, Peto R, Bell J, Zelenika D, Lathrop M, Million Women Study Collaborators: **Gene-environment interactions in 7610 women with breast cancer: prospective evidence from the Million Women Study.** *Lancet* 2010, **375:**2143–2151.
4. Milne RL, Gaudet MM, Spurdle AB, Fasching PA, Couch FJ, Benitez J, Arias Perez JI, Zamora MP, Malats N, Dos Santos SI, Gibson LJ, Fletcher O, Johnson N, Anton-Culver H, Ziogas A, Figueroa J, Brinton L, Sherman ME, Lissowska J, Hopper JL, Dite GS, Apicella C, Southey MC, Sigurdson AJ, Linet MS, Schonfeld SJ, Freedman DM, Mannermaa A, Kosma VM, Kataja V, *et al*: **Assessing interactions between the associations of common genetic susceptibility variants, reproductive history and body mass index with breast cancer risk in the breast cancer association consortium: a combined case–control study.** *Breast Cancer Res* 2010, **12:**R110.
5. Nickels S, Truong T, Hein R, Stevens K, Buck K, Behrens S, Eilber U, Schmidt M, Haberle L, Vrieling A, Gaudet M, Figueroa J, Schoof N, Spurdle AB, Rudolph A, Fasching PA, Hopper JL, Makalic E, Schmidt DF, Southey MC, Beckmann MW, Ekici AB, Fletcher O, Gibson L, Silva Idos S, Peto J, Humphreys MK, Wang J, Cordina-Duverger E, Menegaux F, *et al*: **Evidence of gene-environment interactions between common breast cancer susceptibility loci and established environmental risk factors.** *PLoS Genet* 2013, **9:**e1003284.
6. Campa D, Kaaks R, Le Marchand L, Haiman CA, Travis RC, Berg CD, Buring JE, Chanock SJ, Diver WR, Dostal L, Fournier A, Hankinson SE, Henderson BE, Hoover RN, Isaacs C, Johansson M, Kolonel LN, Kraft P, Lee IM, McCarty CA, Overvad K, Panico S, Peeters PH, Riboli E, Sanchez MJ, Schumacher FR, Skeie G, Stram DO, Thun MJ, Trichopoulos D, *et al*: **Interactions between genetic variants and breast cancer risk factors in the breast and prostate cancer cohort consortium.** *J Natl Cancer Inst* 2011, **103:**1252–1263.
7. Barrdahl M, Canzian F, Joshi AD, Travis RC, Chang-Claude J, Auer PL, Gapstur SM, Gaudet M, Diver WR, Henderson BE, Haiman CA, Schumacher FR, Le Marchand L, Berg CD, Chanock SJ, Hoover RN, Rudolph A, Ziegler RG, Giles GG, Baglietto L, Severi G, Hankinson SE, Lindström S, Willet W, Hunter DJ, Buring JE, Lee IM, Zhang S, Dossus L, Cox DG, *et al*: **Post-GWAS gene-environment interplay in breast cancer: results from the Breast and Prostate Cancer Cohort Consortium and a meta-analysis on 79 000 women.** *Hum Mol Genet* 2014, **23:**5260–5270.
8. Li H, Beeghly-Fadiel A, Wen W, Lu W, Gao YT, Xiang YB, Cai Q, Long J, Shi J, Chen K, Zheng Y, Shu XO, Zheng W: **Gene-environment interactions for breast cancer risk among Chinese women: a report from the Shanghai Breast Cancer Genetics Study.** *Am J Epidemiol* 2013, **177:**161–170.
9. Kraft P, Yen YC, Stram DO, Morrison J, Gauderman WJ: **Exploiting gene-environment interaction to detect genetic associations.** *Hum Hered* 2007, **63:**111–119.
10. Gauderman WJ, Zhang P, Morrison JL, Lewinger JP: **Finding novel genes by testing G x E interactions in a genome-wide association study.** *Genet Epidemiol* 2013, **37:**603–613.
11. Dai JY, Kooperberg C, Leblanc M, Prentice RL: **Two-stage testing procedures with independent filtering for genome-wide gene-environment interaction.** *Biometrika* 2012, **99:**929–944.
12. Schoeps A, Rudolph A, Seibold P, Dunning AM, Milne RL, Bojesen SE, Swerdlow A, Andrulis I, Brenner H, Behrens S, Orr N, Jones M, Ashworth A, Li J, Cramp H, Connley D, Czene K, Darabi H, Chanock SJ, Lissowska J, Figueroa JD, Knight J, Glendon G, Mulligan AM, Dumont M, Severi G, Baglietto L, Olson J, Vachon C, Purrington K, *et al*: **Identification of new genetic susceptibility loci for breast cancer through consideration of gene-environment interactions.** *Genet Epidemiol* 2014, **38:**84–93.
13. Hein R, Beckmann L, Chang-Claude J: **Sample size requirements for indirect association studies of gene-environment interactions (G x E).** *Genet Epidemiol* 2008, **32:**235–245.
14. Dudbridge F, Fletcher O: **Gene-environment dependence creates spurious gene-environment interaction.** *Am J Hum Genet* 2014, **95:**301–307.
15. Johnson N, Walker K, Gibson LJ, Orr N, Folkerd E, Haynes B, Palles C, Coupland B, Schoemaker M, Jones M, Broderick P, Sawyer E, Kerin M, Tomlinson IP, Zvelebil M, Chilcott-Burns S, Tomczyk K, Simpson G, Williamson J, Hillier SG, Ross G, Houlston RS, Swerdlow A, Ashworth A, Dowsett M, Peto J, Dos Santos SI, Fletcher O: **CYP3A variation, premenopausal estrone levels, and breast cancer risk.** *J Natl Cancer Inst* 2012, **104:**657–669.
16. Johnson N, Dudbridge F, Orr N, Gibson L, Jones ME, Schoemaker MJ, Folkerd EJ, Haynes BP, Hopper JL, Southey MC, Dite GS, Apicella C, Schmidt MK, Broeks A, Van T, Veer LJ, Atsma F, Muir K, Lophatananon A, Fasching PA, Beckmann MW, Ekici AB, Renner SP, Sawyer E, Tomlinson I, Kerin M, Miller N, Burwinkel B, Marme F, Schneeweiss A, *et al*: **Genetic variation at CYP3A is associated with age at menarche and breast cancer risk: a case–control study.** *Breast Cancer Res* 2014, **16:**R51.
17. Swerdlow AJ, Jones ME, Schoemaker MJ, Hemming J, Thomas D, Williamson J, Ashworth A: **The Breakthrough Generations Study: design of a long-term UK cohort study to investigate breast cancer aetiology.** *Br J Cancer* 2011, **105:**911–917.
18. Easton DF, Pooley KA, Dunning AM, Pharoah PD, Thompson D, Ballinger DG, Struewing JP, Morrison J, Field H, Luben R, Wareham N, Ahmed S, Healey CS, Bowman R, SEARCH collaborators, Meyer KB, Haiman CA, Kolonel LK, Henderson BE, Le Marchand L, Brennan P, Sangrajrang S, Gaborieau V, Odefrey F, Shen CY, Wu PE, Wang HC, Eccles D, Evans DG, Peto J, *et al*: **Genome-wide association study identifies novel breast cancer susceptibility loci.** *Nature* 2007, **447:**1087–1093.
19. Hunter DJ, Kraft P, Jacobs KB, Cox DG, Yeager M, Hankinson SE, Wacholder S, Wang Z, Welch R, Hutchinson A, Wang J, Yu K, Chatterjee N, Orr N, Willett WC, Colditz GA, Ziegler RG, Berg CD, Buys SS, McCarty CA, Feigelson HS, Calle EE, Thun MJ, Hayes RB, Tucker M, Gerhard DS, Fraumeni JF Jr, Hoover RN, Thomas G, Chanock SJ: **A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer.** *Nat Genet* 2007, **39:**870–874.
20. Stacey SN, Manolescu A, Sulem P, Rafnar T, Gudmundsson J, Gudjonsson SA, Masson G, Jakobsdottir M, Thorlacius S, Helgason A, Aben KK, Strobbe LJ, Albers-Akkers MT, Swinkels DW, Henderson BE, Kolonel LN, Le Marchand L, Millastre E, Andres R, Godino J, Garcia-Prats MD, Polo E, Tres A, Mouy M, Saemundsdottir J, Backman VM, Gudmundsson L, Kristjansson K, Bergthorsson JT, Kostic J, *et al*: **Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor-positive breast cancer.** *Nat Genet* 2007, **39:**865–869.

doi:10.1186/s12916-014-0195-1
Cite this article as: Fletcher and Dudbridge: **Candidate gene-environment interactions in breast cancer.** *BMC Medicine* 2014 **12**:195.